REMARKS/ARGUMENTS

Claims 13, 15, 18-20 and 27-31 are active in this application.

Support for the amendment to Claim 13 is found in the sequence listing, Claim 14, and the specification on page 5, line 18 and line 24.

The specification is amended to include a brief description of the drawings and Sequence Identifiers (SEQ ID NO:) where appropriate. Support for the description of the drawings is found in the specification on pages 8-15.

No new matter is added by these amendments.

The rejection of Claims 13-20 under 35 U.S.C. § 112, first paragraph ("written description") is respectfully traversed.

The polypeptide as defined in the pending claims is characterized by (1) being an insecticidal polypeptide; (2) obtained from a legume seed; (3) sequence of formula I where C represents a cysteine residue, X_1 represents a dipeptide, X_2 represents a tripeptide, X_3 represents a heptapeptide, X_4 represents a tetrapeptide, X_5 represents an amino acid, X_6 represents a nonapeptide, and X_7 represents a pentapeptide; and (4) the sequence has at least 60% identity with SEQ ID NO:6 or SEQ ID NO:7.

The fundamental question to be asked when assessing whether a set of claims is adequately described by the specification is: whether the specification describes "the claimed invention in sufficient detail that one skilled in the art can reasonably conclude that the inventor had possession of the claimed invention." [1] The specification unquestionably describes the polypeptides used in the claim method demonstrating possession of the claimed invention.

[1] *Vas-Cath, Inc. v. Mahurkar*, 935 F.2d 1555, 1563, 19 USPQ2d 111, 1116, (Fed. Cir. 1991).

The Examiner contends that "the specification does not describe a representative number of species of insecticidal proteins that meet the structural limitations of the claims." (page 5 of the Official Action). Applicants respectfully disagree. First and foremost, the Examiner is narrowly focused on three specific proteins described, for example, in Figure 7, i.e., TP protein, PA1b pea albumin and leginsulin. However, the specification is not so limited in its description. Applicants have described a polypeptide having a specific formula, with a limited set of substitutions within certain defined locations. Each of the polypeptides resulting from those substitutions of X (1-7) in formula (I) is one representative specie. Continuing with further substitutions at the X positions in formula (I) yields a second specie and so on until all of the species are envisioned. It is implausible to argue that one could not appreciate all of the species described in the specification and encompassed by the polypeptide in the claimed method. Therefore, the Examiner's focus on three specific amino acids rather than the polypeptide formula itself is improper.

Furthermore, Applicants have presented an alignment in Figure 7, which as described on page, 3 lines 11-27 demonstrated 6 conserved cysteine residues, which are required in the polypeptide of formula (I) as claimed.

Accordingly, withdrawal of this ground of rejection is requested.

The rejection of claims 13-20 under 35 U.S.C. § 112, first paragraph ("enablement") is respectfully traversed.

This rejection is believed to be overcome at least in part, based on the definition of the polypeptide as obtained from the seeds of a legume. In addition, as discussed above concerning the written description rejection, the Applicants have described numerous species of polypeptides. As on basis for asserting this rejection, the Examiner contends that there is no guidance "with respect to the specific amino acid structural elements that would be

retained by insecticidally active forms of these variants that would further function to protect plants from insects." (Page 9 of the Official Action). This is not correct. As noted above, described on page 3 of the application and defined in the polypeptide claimed, the polypeptides used in the claimed methods all have at least the seven (7) conserved cysteine amino acids.

As a further rationale for this rejection, the Examiner indicates that the present invention would require undue experimentation to identify polypeptides that function to protect plants from insects. However, the Examiner appears to be confusing the burden of "undue experimentation" with the "amount of work."

The Examiner's attention is drawn to In re Wands, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988), which states: "Time and difficulty of experiments are not determinative if they are merely routine." (see MPEP §2164.06). Again citing In re Wands, MPEP §2164.06 states: "The test is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed."

The specification on page 5 describes obtaining polypeptides from the seeds of legumes. The specification, in Example 2, describes how to isolate such polypeptides and in Example 1 how to confirm their insecticidal activity. In fact, insecticidal polypeptides within the scope of the insecticidal polypeptide claimed were obtained and described in the attached publication of Louis et al (*Plant Science* 167(2004):705-714)—see Figure 4 on page 710 and the Abstract on page 705.

Therefore, the claimed invention is enabled by the specification as originally filed and as such withdrawal of this ground of rejection is requested.

The rejection of Claim 13 under 35 U.S.C. § 103(a) in view of Iizuka (U.S. patent no. 5,516,514) and Raikhel (U.S. patent no. 5,276,269) is respectfully traversed. Claim 13 has been amended, in part, to include the limitations of Claim 14 which was not rejected by this combination of publications. Furthermore, the polypeptides in each of the cited publications are not the same as the polypeptide defined in the present claims. Therefore, the claimed invention would not have been obvious in view of these two publications.

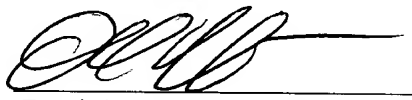
Withdrawal of this ground of rejection is requested.

The rejection of Claim 13, and claims dependent thereon, under 35 U.S.C. § 112, second paragraph is obviated by the amendment submitted herein.

Applicants request allowance of all pending claims.

Respectfully submitted,

OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.
Norman F. Oblon



Daniel J. Pereira, Ph.D.
Attorney of Record
Registration No. 45,518

Customer Number
22850

Tel: (703) 413-3000
Fax: (703) 413 -2220
(OSMMN 06/04)

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Plant Science 167 (2004) 705–714
PLANT SCIENCE
www.elsevier.com/locate/plantsci

Molecular and biological screening for insect-toxic seed albumins from four legume species

Sandrine Louis [a], Bernard Delobel [a], Frédéric Gressent [a], Isabelle Rahioui [a], Laurence Quillien [b,1], Agnès Vallier [a], Yvan Rahbé [a,*]

[a] *UMR INRA-INSA de Lyon BF2I (Biologie Fonctionnelle Insectes et Interactions), Bat. Louis-Pasteur, 20 Avenue A. Einstein, 69621 Villeurbanne Cedex, France*
[b] *INRA URPVI (Unité de Recherches sur les Protéines Végétales et leurs Interactions), rue de la Géraudière BP 71627, 44316 Nantes Cedex 03, France*

Received 17 September 2003; received in revised form 16 April 2004; accepted 20 April 2004

Available online 18 May 2004

Abstract

Pea albumin 1b (PA1b) is a small sulphur-rich peptide from pea seeds, also called leginsulin due to the binding properties of its soybean orthologue. Its insecticidal properties were discovered more recently. By using a combination of molecular, biochemical and specific insect bioassays, we characterised new genes and their products from the seeds of four legume species. Two species (*Glycine max* and *Phaseolus vulgaris*) display most of the characteristics of the *Pisum sativum* type: homologous genes and predicted toxic hydrophobic peptides with similar post-translational processing. The third species (*Medicago truncatula*) possesses homologous genes and high insecticidal activity, but no specific biochemical detection of the peptide products was obtained, indicating possible variant post-translational processing. Our combined approach appears to be efficient for a broad study of A1b within legumes.

Keywords: Albumins 1b; PA1b; Insect; Weevil; Plant defence; Toxin; French bean; Barrel medic

1. Introduction

Insect–plant interactions are generally governed by complex sets of physiological and chemical determinants controlling either host plant acceptation (choice of plant) or host plant adequacy (success on the plant). Defensive mechanisms in response to phytophagous invertebrates are sometimes grouped as antixenosis (counteracting acceptation), antibiosis (counteracting physiological adequacy) and plant tolerance [1]. An effective combination of these factors leads to plant resistance, often based on complex genetics [2,3], although simple traits have also been identified as formal resistance genes in some crop species, especially against small parasitic-like insects [4]. Recently, a simple genetic system was identified in the pea, the seeds of which are protected from cereal weevils (*Sitophilus* spp.) by a small polypeptide which kills the non-host pest after a few days of seed consumption [5]. In *Sitophilus oryzae*, some strains were found to harbour a single recessive gene responsible for full immunity to this peptide [6]. This molecule was previously known and cloned as the sulphur-rich Pea albumin 1b (PA1b) [7], with no other known function than that of sulphur-storage. It is a small hydrophobic knotted peptide of 37 amino acids with three disulphide bridges, highly stable to thermal and protease inactivation, even within the insect gut [5,8]. Its structure has been recently elucidated [9,10],

Abbreviations: cv, cultivar (cultivated plant genotype); ESI–MS, electrospray ionisation mass spectrometry; MALDI–TOF MS, matrix assisted laser desorption ionisation, time-of-flight detection mass spectrometry; MeOH/MeOH60/H_2O5/H_2O8, methanolic fraction (respectively 60% aqueous methanol/acidic water/basic water extracts, see text for details); EST, expressed sequence tag; PA1b/a, pea albumin 1 subunit b (respectively a); PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecyl sulfate; UTR, untranslated region; TC, tentative consensus sequence cluster

* Corresponding author. Tel.: +33-4-72438476/72438356; fax: +33-4-72438534.
E-mail addresses: louis.sandrine@jouy.inra.fr (S. Louis), delobel.bernard@jouy.inra.fr (B. Delobel), gressent.frederic@jouy.inra.fr (F. Gressent), rahioui.isabelle@jouy.inra.fr (I. Rahioui), laurence.quillien@nantes.inra.fr (L. Quillien), vallier.agnes@jouy.inra.fr (A. Vallier), rahbe@jouy.inra.fr (Y. Rahbé).
[1] Tel.: +33-2-40675036; fax: +33-2-40675025.

0168-9452/$ – see front matter © 2004 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2004.04.018

and it has been shown to interact with a microsomal high affinity binding-site identified in susceptible insects but absent in resistant strains [11].

The albumin 1 gene identified in pea encodes a preproprotein with a signal peptide, the toxic 37-residue PA1b, a small linker peptide, and another sulphur-rich 53-residue polypeptide called PA1a, of unknown function [7]. A1b was also later identified from soybean seeds as a natural endogenous ligand to an insulin-binding globulin, hence the name leginsulin [12]. Although a few more related sequence data have been made available since [13], the only other work on leginsulin appeared recently and was concerned only with its 7S globulin-binding properties [10,14].

Following the identification of a major defensive function for this peptide, and as a prerequisite to a comprehensive survey of this gene family in the Fabaceae, our aim was to validate a simultaneous molecular, biochemical and biological characterisation of this toxin from legume seeds, mainly through a combination of genomic PCR, non-polar peptide analysis and PA1b-specific bioassays. Using this combined approach on *Pisum sativum*, *Glycine max* and two other species as yet uncharacterised for this peptide family (*Phaseolus vulgaris* and *Medicago truncatula*), we cloned homologous genomic sequences, identified most of the resulting product peptides and quantified the corresponding biological activities from all seed extracts. Four PA1b alleles/loci were identified from the pea genotype Frisson (among a potential of ten, from peptide analysis). In spite of already extensive expressed sequence tag (EST) analysis in *M. truncatula*, the two cloned PA1b genes from this species were new. Through analysis of the EST databases from this species, we also identified the A1 (albumins 1) as a small multigenic family, with previously unrelated members, which are expressed outside the seeds.

2. Material and methods

2.1. Insects

Rice weevils (*S. oryzae*, Coleoptera Curculionidae) were reared on wheat seeds at 27.5 °C and 70% RH. Two strains were used, differing in their genetic ability to thrive on pea seeds and resist the toxic activity of pea albumin PA1b: a control susceptible strain "Benin" (S) and a fully resistant strain "China" (R) harbouring the recessive pea-resistance allele [6].

2.2. Plant material and peptide extractions

We used seeds from pea *P. sativum* L. cultivar (cv) Frisson (tribe Vicieae; gift of G. Duc INRA Dijon), soybean *G. max* L. cv Paoki (tribe Phaseoleae; gift from P. Sartre, INRA Montpellier), bean *P. vulgaris* L. cv Contender (tribe Phaseoleae; commercially available), and of the model legume *M. truncatula* Gaertner cv Salernes (tribe Trifolieae, gift from J.M. Prosperi, INRA Montpellier). Seeds (200–1000 g) were crushed in a Warring blender and sieved through a 0.4 mm mesh to separate the cuticles from the flour. A few seeds were planted in a greenhouse to obtain young leaf material subsequently used for DNA extraction. Flours were submitted to successive extractions aimed at fractionating peptides from either apolar compounds or proteins denatured by the serial procedure (100 g + 1 L of solvent, overnight stirring, filtering, vacuum drying in a Buchi Rotavapor®). The successive solvents used were pentane, methanol 100% (resulting in fraction labelled MeOH), methanol 60% in water (MeOH60), water pH 5 (H_2O5) and water pH 8 (H_2O8). Soluble fractions, except for pentane, as well as the final residue (Res) were bioassayed for S and R weevil toxicity. To further purify strict PA1b homologues, MeOH60 extracts were solubilised in acetone 80%, placed at −20 °C for 45 min and centrifuged for 20 min at 12 000 × *g* and at 4 °C [11]. When needed, individual peptides were further purified by RP-HPLC as described below.

2.3. Bioassays

Insects used for bioassays were adults aged 2–3 weeks, collected from experimental 1-week cohorts and deposited in batches of 30 individuals (for each S and R strain) on food pellets incorporating tested seed flours, or flour fractions, in a whole-wheat based diet. Whole flours were tested over a 5–60% range ((w/w) in wheat). Concentrations of the fractionated material were always given as equivalent (%) of their relative abundance in their original legume flour (total meal equivalent (%), TME). Bioactivity was evaluated by scoring daily insect survival during the first 2 weeks of contact with the test food (27.5 °C, 70% RH), and by standard survival analysis, followed by a LT50 calculation (lethal time 50%, or median life duration; Statview software, actuarial analysis and associated non-parametric statistics).

2.4. Electrophoresis, antibodies and Western-blotting

Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was conducted on 16% gels (A-3574, Sigma, France), and revealed with Coomassie Brilliant Blue G 250 (0.5 g/L). Loaded proteins were quantified by a Bradford assay [15].

HPLC-purified PA1b was conjugated N-terminally to ovalbumin before using it to immunise rabbits [16]. For coupling, PA1b was added to a 20 mg/ml solution of ovalbumin in phosphate buffer saline (PBS, 10 mM NaH_2PO_4, 150 mM NaCl, pH 7.4) at a molar ratio protein/peptide of 1/40. An equal volume of a fresh glutaraldehyde solution was added drop wise to the mixture under constant stirring (final concentration: 1% (w/v)). After 1 h at 40 °C, the reaction was stopped by adding sodium borohydride (10 mg/ml). Dialysis against PBS was performed to remove free PA1b and linkers.

Rabbit polyclonal antibodies were obtained by subcutaneous immunisation, initially with complete Freund's adjuvant, then every 15 days with incomplete adjuvant. Bleeding was performed after five injections. The specificity of antibodies was tested by immunoblotting.

Western-blots were performed by a liquid blotting of gels on Protan BA 83 nitrocellulose membranes (Schleicher and Schuell, France) in a modified methanol-free buffer (25 mM Tris, 192 mM glycine) for 45 min at 250 mA (Bio-Rad trans-blot® cell). Membranes were then blocked for 16 h in 5% skimmed milk in Tris buffer saline (TBS: 50 mM Tris, 200 mM NaCl, pH 7.4). Antibodies were then used at 1/500 dilution for 1 h, and the membranes were subsequently rinsed in Tween 20 PBS (0.05% detergent in PBS). Peroxidase-coupled goat anti-rabbit secondary antibody (170-6515, Bio-Rad, France) was then used with HRP colour development reagent according to the manufacturer's instructions (170-6534, Bio-Rad, France).

2.5. HPLC

Reverse-phase HPLCs were performed on a Nucleosil® 300 C18 column (250 mm × 4.6 mm; 5 μm particle size, 300 Å porosity). Proteins were eluted at 1 mL/min with a 22 min gradient from 20 to 60% acetonitrile in water (0.1% TFA), and monitored by UV diode-array detection between 210 and 350 nm.

2.6. Mass spectrometry

The 200 μg of target plant fractions containing toxicity and potential A1b peptide (mainly MeOH60 fractions, see results) were submitted to mass spectrometry on a Voyager DE-PRO spectrometer (PerSeptive Biosystems, Farmingham, MA, USA). Positive ion mass spectra were recorded in the linear mode of this time-of-flight MALDI mass spectrometer. All mass spectra were externally calibrated with a calibration kit (Pep Mix 2, LaserBio Labs, Sophia Antipolis France), allowing a mass accuracy ±0.05%. Samples were mixed with the sinapinic acid matrix (3.5-dimethoxy-4-hydroxy-cinnamic acid, LaserBio Labs) at a ratio ranging from 1:1 to 1:10, then spotted onto the target, dried, and submitted to MALDI–TOF analysis.

For pea aqueous methanolic extracts, HPLC was performed on the solvent extract, as mentioned above, and A1b peaks (retention times: 16–20 min) were collected and analysed by ESI–MS on a triple quadrupole mass spectrometer API III+ SCIEX (Thornhill, Ont., Canada).

2.7. Binding assays

A purified isoform of PA1b (PsaA1b005, MW: 3741 Da by mass spectrometry) was labelled with ^{125}I to a specific radioactivity of ca. 1000 Ci/mmol and used in binding assays as described [11], with target membrane proteins extracted from the *S. oryzae* susceptible strain Bénin. For competition data analysis, and for each plant, comparative quantification of binding inhibition was expressed as the mass of total meal equivalent present in the MeOH60 fraction, and needed to inhibit 50% of the radiolabelled ligand binding. As a negative control, wheat meal was extracted, as described, and its MeOH60 fraction was assayed.

2.8. DNA extraction and degenerate primer genomic PCR

DNA was extracted from fresh leaves using a CTAB based protocol [17] adding, during grinding, 15% (w/w) polyvinylpolypyrrolidone (MW: 40 000) to eliminate any phenolics [18]. PCR amplifications (35 cycles: 94 °C (30 s), 50 °C (45 s), and 72 °C (45 s)) were realised on 10 ng of genomic DNA with for/rev pairs (final concentration of 3 μM each) of the following degenerated primers designed from sequences of soybean and pea, either in the A1b (For 1) or the A1a peptides (Rev 1 and Rev 3): For 1: 5′-TGYTCICCITTYGARRTICCICCITG-3′, Rev1: 5′-CRAARCACCAICCRTAITCIATRTM-3′, Rev3: 5′-SRCAIARRTTIGGRTSYTCITC-3′.

PCR amplified fragments were purified, ligated to pMos-Blue vectors and used to transform *E. coli* NM 522 electrocompetent cells. Recombinant plasmids were purified and inserts sequenced (Genome Express, Grenoble, France). Four clones were sequenced from each first round of PCR in each plant species to select potential variants (loci/alleles).

For pea (cv. Frisson), RNA was extracted from mid-growth seeds [19], and cDNAs were obtained with the SuperScript RT-PCR kit (Invitrogen). PCR was run on either leaf genomic or seed cDNAs with primers For PA1b 5′-ATCAAACAATGGCTTCCGTTAAA-3′ and Rev PA1b 5′-TCGAAATTAAGCAGTGGAAACAC-3′ (30 cycles: 94 °C (30 s), 53.5 °C (45 s), and 72 °C (45 s)). PCR products were cloned into a pCR 2.1 plasmid, and transformed into *E. coli* Top10 cells. Twelve cloned inserts were double-sequenced (Genome Express, Grenoble France) to yield pea genomic and cDNA sequences.

2.9. Genome walking for cloning of A1b 5′ ends and sequence analysis

Gene walking was carried out using the Universal GenomeWalker™ kit (Clontech, USA) with primers designed from specific parts of the sequences obtained in the first round of degenerated PCR. Each genomic DNA was digested by four restriction enzymes (DraI, EcoRV, StuI, PvuII). Adaptors were ligated to the restriction fragments and two successive nested PCR amplifications were performed on the products of ligation, following the manufacturer's instructions. The major PCR product obtained over 600 bp from the four restriction libraries was cloned, and two clones were double-sequenced for each primer pair. Sequences were assembled and analysed with the MacMolly software, checked for sequencing errors, for adequacy to first-round PCR results, annotated for cds,

intron position and signal peptide positions (using SignalP, Netgene2, Netstart at http://www.cbs.dtu.dk/, and manual refinement with published pea and soybean genomic sequences as templates), and finally submitted to EMBL with the following accessions numbers: AJ574789–AJ574796.

3. Results

3.1. Whole meal bioassays

The results of whole seed toxicity to the rice weevils are reported in Fig. 1, showing that all four plants were highly toxic, to a standard (susceptible) strain of *S. oryzae* at meal concentrations higher than 20% (Fig. 1A). On day 4, the differences in total weevil mortality between plant species and meal concentrations were the most pronounced. In *P. vulgaris* and *M. truncatula*, the dose–response curve shapes were indicative of deterrent factors (antixenosis), inhibiting food and thus toxin uptake at high doses (less mortality at 80% than at 40% of plant meal); fasting insects did not die in these assay conditions (not shown), which was exemplified by the absence of mortality of R strain on the *P. vulgaris* meal, even at high doses. Given the genetic background, the comparison of susceptible and resistant strains indicates that most of the observed mortality was caused by albumin 1b-like components (differential S/R toxicity), except in *M. truncatula* for which high mortality of R-strain occurred at intermediate doses (Fig. 1B).

3.2. Bioassays of seed peptide fractions

To identify the fractions containing potential A1b homologues, comparisons were carried out between the bioassay toxicity results on S and R strains (the lethal time 50 for each fraction and strain pair is shown in Table 1), based on survival analysis of data similar to those shown for the MeOH60 fractions (Fig. 2). In pea, soybean and bean, the highest A1b-related toxicity (mortality of the S strain weevils and no mortality of the R strain insects) was present in the MeOH60 fractions, although other fractions still carried variously distributed amounts of R/S differential toxic

Table 1
Toxicity of seed fractions to weevil strains susceptible (S) and resistant (R) to pea albumin PA1b

Plant species	Seed fraction	S strain (days)	R strain (days)	p (S vs. R)
P. sativum	MeOH	Not lethal	Not lethal	–
	MeOH60	4.5 ± 0.2	Not lethal	–
	H_2O5	5.7 ± 0.4	Not lethal	–
	H_2O8	5.0 ± 0.4	Not lethal	–
	Residue	5.4 ± 0.3	Not lethal	–
G. max	MeOH	7.3 ± 0.9	Not lethal	–
	MeOH60	6.5 ± 0.9	Not lethal	–
	H_2O5	Not lethal	Not lethal	–
	H_2O8	10.8 ± 0.6	Not lethal	–
	Residue	8.0 ± 1.3	Not lethal	–
P. vulgaris	MeOH	Not lethal	Not lethal	–
	MeOH60	6.3 ± 0.3	Not lethal	–
	H_2O5	Not lethal	Not lethal	–
	H_2O8	8.8 ± 0.9	Not lethal	–
	Residue	10.6 ± 0.2	Not lethal	–
M. truncatula	MeOH	3.9 ± 0.4	8.8 ± 0.6	<0.0001
	MeOH60	3.5 ± 0.3	3.6 ± 0.2	0.65
	H_2O5	4.5 ± 0.2	6.5 ± 0.3	<0.0001
	H_2O8	4.7 ± 0.3	5.3 ± 0.2	0.008
	Residue	4.3 ± 0.4	5.2 ± 0.2	0.002

Median survival time (±S.E.) on diets incorporating 100% seed equivalent of each fraction. Test of actuarial survival analysis, compared between both strains, is reported in the last column (*P* value of Breslow–Gehan–Wilcoxon statistics).

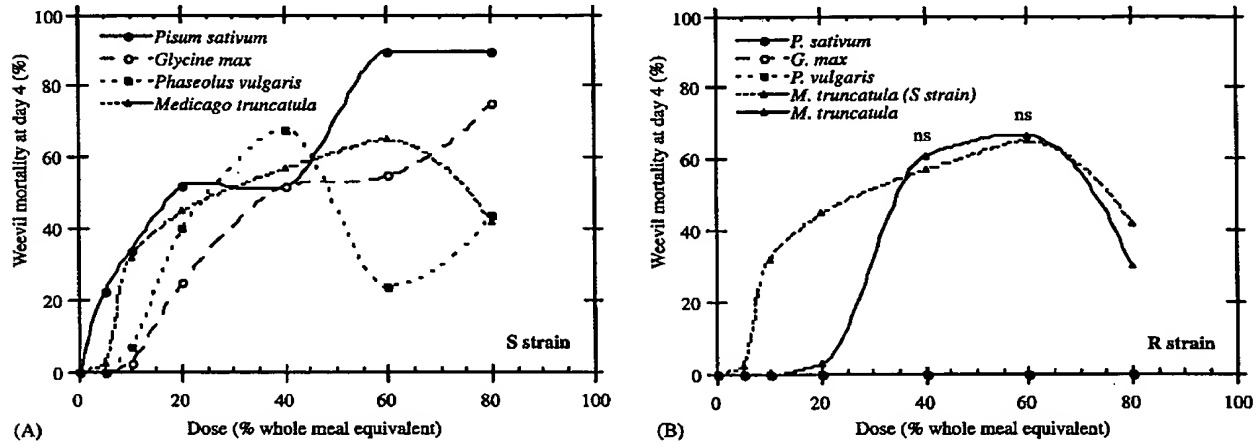


Fig. 1. Acute toxicity to weevils of seed flour diet from four legume species. Dose response curves (% legume in wheat diet) on two weevil strains, one control susceptible S strain (A) and one resistant R strain to the PA1b toxin (B); data from S strain on *M. truncatula* are reported in (B) for comparison (ns: non-significant test in R vs. S, using survival analysis).

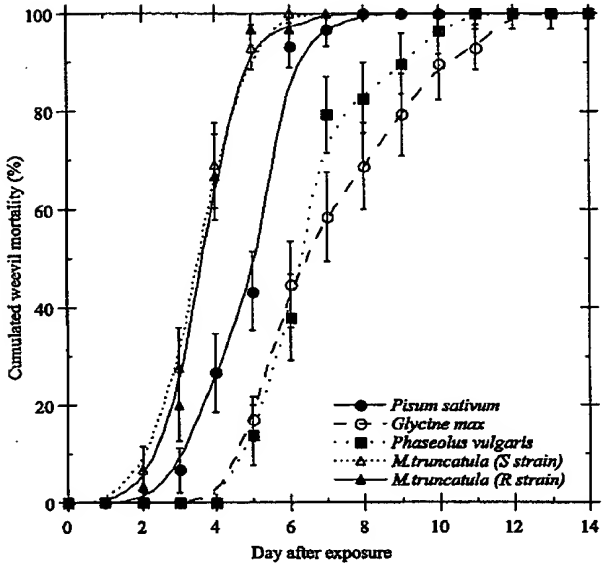


Fig. 2. Time-course of weevil mortality on MeOH60 fractions from the four legume species (dosed at 100% original seed equivalent, in wheat). All data from susceptible S strain, plus R strain for *M. truncatula* extracts. Standard error bars are reported from survival analysis estimates.

activity, possibly due to partial cross-contamination with PA1b peptides (Table 2). Nevertheless, cross-contamination was probably restricted to some molecular species only (see Section 3.5) as it was not detectable biologically in the acidic water extract from *G. max* and *P. vulgaris* (Table 1). Furthermore, the persistent differential activity observed in the insoluble fraction of these species (residue; Table 1) is most easily interpreted as a carryover of bound forms of A1b-like peptides, which were detected *in vivo* in membrane/cell-wall fractions of soybean [10]. In *M. truncatula*, however, the differential activity observed on whole meals was mainly recovered in the methanolic fraction (and marginally in the H_2O5 extract), and the MeOH60 extract was characterised in that it carried an activity that was highly toxic towards the PA1b-resistant weevil strain (same toxicity to both strains).

3.3. A1b ligand binding competition with hydrophobic peptide fractions

To further assess the presence or absence of PA1b homologues in fractions selected after bioassays (all MeOH60 extracts, plus all *M. truncatula* fractions) a ligand binding competition assay was carried out, between these extracts and radiolabelled pea toxin binding to its high affinity

Table 2
Mass spectrometry peptide matches on selected solvent extracts from the four plant species analysed

Plant species (genotype)	Theoretical masses[a] (target sequences)	Fraction	Observed masses matching[b] (relative intensity)	Matched intensity[d] (%)
Glycine max (cv. Paoki)	3876.5 (GmaA1b005), 3819.5 (c-1)	MeOH60	3817 (25), 3874 (100 %), 3890 (15), 3928 (16), 7858 (18)	72
Medicago truncatula, (cv. Salernes)	3870.4 (MtrA1b006), 3858.4 (MtrA1b007)	MeOH	No mass >2000[c]	–
		MeOH60	2593 (54), 2767 (24), 3443 (23), 4375 (43), 4650 (12), 5114 (23), 5186 (100%), 6887 (19)	0%
Phaseolus vulgaris, (cv. Contender)	3994.6 (PvuA1b001), 3937.6 (c-1)	MeOH60	3764 (68), 3809 (100%), 3820 (89), 3822 (83), 3866 (80), 3868 (75), 3874 (23), 3935 (62), 3993 (25) *2718, 4481, 5439*	14
Pisum sativum (cv. Frisson)	3742.4 (PsaA1b011), 3685.3 (c-1)	MeOH	traces of *3745, 3760,3791, 3808 3819*[c]	–
	3742.4 (PsaA1b014), 3789.5 (PsaA1b015), 3732.4 (c-1), 3790.4 (PsaA1b012), 3733.4 (c-1), 5908.4 (PsaA1a014), 5923.5 (PsaA1a015)	MeOH60	3731 (16), 3742 (35), 3758 (14), 3789 (86), 3806 (17), 3816 (15), 5541 (14), 5740 (31), 5907 (56), 5921 (53), 5935 (53), 5968 (12), 5983 (100%), 5998 (14), 6014 (20), 6028 (15) *3685, 3774, 3843*[c]	67
	5971.5 (PsaA1a011), 5984.5 (PsaA1a012)	H_2O pH 5	3729 (23), 3743 (24), 3790 (30), 3818 (30), 5005 (10), 5741 (28), 5908 (56), 5923 (50), 5936 (53), 5968 (10), 5984 (100), 6015 (18), 6029 (23), 6070 (13), 6086 (13), 6099 (13), 6114 (11), 6130 (12), 6146 (22), 6191 (17), 6209 (10), 6221 (11), 6809 (18), 7850 (27)	49

[a] Expressed as MH+ (from nucleic sequences obtained, average mass minus six for full cysteine bridge bonding); when a hit was discovered with a truncated sequence, the theoretical mass of the truncated sequence is reported (c—terminal truncation, based on a canonical pea-type pre- and propeptide processing).
[b] Matching was scored positive when a precision better than 0.05% was obtained with the theoretical mass (linear detection mode); observed masses are rounded to the closest mass unit.
[c] We only reported peaks with *m/z* > 2000, and of intensity exceeding 10% of the highest observed peak (other significant but minor peaks are listed in italics).
[d] Expressed as percentage of matched peak intensities vs. total intensity of major peaks (as defined in c, using peak heights).

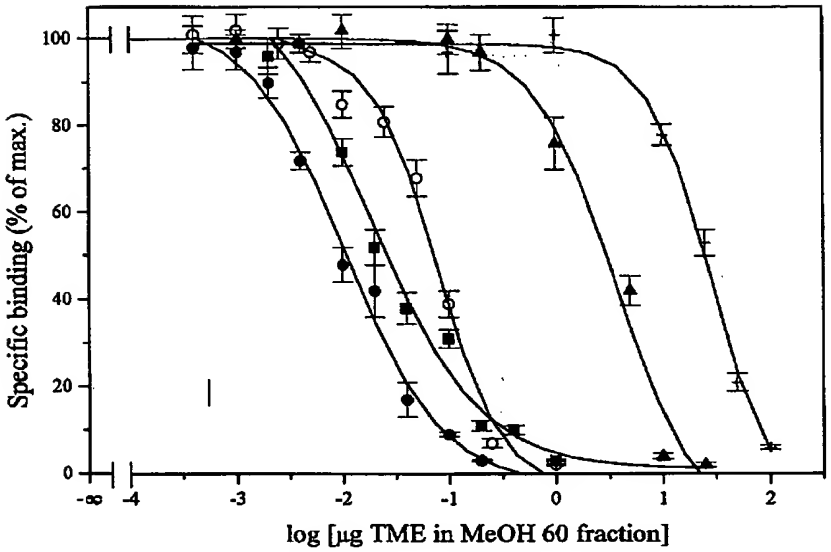


Fig. 3. Competitive inhibition of ^{125}I-PA1b binding to its binding-protein by MeOH60 extracts from the four test plant species, plus wheat control. *X*-axis: dose of extract (total meal equivalents, TME, present in the aqueous methanolic extract). Source plants: (●) *Pisum sativum*, (■) *Phaseolus vulgaris*, (○) *Glycine max*, (+) *Triticum aestivum* and (▲) *Medicago truncatula*. Experimental standard error bars are reported.

binding site (K_d of 6 nM), on susceptible insect membrane preparations. The results show that all MeOH60 extracts, except that of *M. truncatula*, were able to strongly inhibit the ^{125}I-PA1b specific binding to the insect membrane proteins (Fig. 3). For *M. truncatula*, the binding inhibition was only 10 times more than that of the wheat negative control. Other negative controls included the methanolic and acidic water fractions from wheat (not shown), and the *M. truncatula* water fractions, which did not show any binding inhibition activity. For the three clearly positive species, 50% inhibition of ^{125}I-PA1b binding occurred for 10 (±0.5), 18 (±2) and 79 (±3) µg total-meal-equivalent in the MeOH60 extracts (from *P. sativum*, *P. vulgaris* and *G. max*, respectively). These values are about two to three orders of magnitude smaller than those obtained with the negative control, and are in good ranking agreement with the corresponding bioassay figures (Fig. 2).

3.4. Cloned albumin 1b genes from the four plant species

Complete A1b sequences were obtained for the four species, using PCR with degenerate primers and gene walking for 5′completion, plus cDNA cloning for *P. sativum*: two sequences were obtained from *M. truncatula*, encompassing the whole genes and significant 5′UTR (untranslated region) regions (AJ574789 and AJ574790), four from *P. sativum* (AJ574793-6), one from *G. max* (AJ574791), and one for *P. vulgaris* (AJ574792) The corresponding protein sequences are shown in Fig. 4. One incomplete sequence

```
                                <---------- leader ------> <----------------- PA1b ------------>        <-- PA1a' -->
PsaA1b005-3741.39 (Psa, ref. 5)                            ASCNGV---CSPFEMPPCGTSA-CRCIPVGLVI-GYCRNPSG
Q40999 (Psa, pa1 M81864)                  MASVKLA-SLIVLFATLGMFLTKNVGA ASCNGV---CSPFEMPPCGSSA-CRCIPVGLLI-GYCRNPSGV--FLKGNDEHPNLC/...
PsaA1b011 (Psa, AJ574793)                 MASVKLA-SLIVLFATLGMFLTKNVGA ASCNGV---CSPFEMPPCGTSA-CRCIPVGLVI-GYCRNPSGV--FLRTNDEHPNLC/...
PsaA1b012 (Psa, AJ574794)                 MASVKLA-SLIVLFATLGMFLTKNVGA ASCNGV---CSPFEMPPCGTSA-CRCIPVGLFI-GYCRNPSGV--FLKANDEHPNLC/...
PsaA1b014 (Psa, AJ574795)                 MASVKLA-SLIVLFATLGMFLTKNVGA ASCNGV---CSPFEMPPCGSSA-CRCIPVGLLI-GYCRNPSGV--FLKGNDEHPNLC/...
PsaA1b015 (Psa, AJ574796)                 MASVKLA-SLIVLFATLGMFLTKNVGA ISCNGV---CSPFDIPPCGSPL-CRCIPAGLVI-GNCRNPYGV--FLRTNDEHPNLC/...
BE661090 (Gma, leginsulin EST)            MAYARLAPMAVFLLATSTIMFPTKIEA ADCNGA---CSPFEMPPCRSRD-CRCVPIGLVA-GFCIHPTGLSSVAKMIDEHPNLC/...
GmaA1b005 (Gma, AJ574791)                   >ARLAPMAVFLLATSTIMFPTKIEA ADCNGA---CSPFEMPPCRSRD-CRCVPIGLVA-GFCIHPTGLSSVAKMIDEHPNLC/...
Q39837 (Gma, leginsulin AJ223037)                MAVFLLATSTIMFPTKIEA ADCNGA---CSPFEVPPCRSRD-CRCVPIGLFV-GFCIHPTGLSSVAKMIDEHPNLC/...
Q9ZQX0 (Gso, leginsulin AJ011935)                MAVFLLATSTIMFPTKIEA ADCNGA---CSPFEVPPCRSSD-CRCVPIGLFV-GFCIHPTGLSSVAKMVDEHPNLC/...
PvuA1b041 (Pvu, unpublished)                                          /---CSPFEMPPCGSSD-CRCVPYGLFV-GSCIHPTGLSAAAKMIDEHPNLC/...
PvuA1b001 (Pvu, AJ574792)                 MANVRVAPLALFLLATSIMFPMKKTEA VVCSGV---CSPFERPPCGSTRDCRCIPYGLFI-GACTYPSGLSSVAKTIDEHPNIC/...
MtrA1b006 (Mtr, AJ574789)                 MAYIRFAHLVVFLLAAFSLVPTKKVGA TDCSGA---CSPFEMPPCRSSD-CRCIPIGLVA-GYCTYPSSP-TVMKMVEEHPNLC/...
MtrA1b007 (Mtr, AJ574790)                 MAYLRLAHLVVFLHATFSLIFPMMKAA EDCSGI---CSPFEMPPCPSSS-CRCIPVILIG-GNYVDPSSP-TITKMVEKHANLC/...
AJ389043 (Mtr, EST)                       MTYVKLAILAVLHLTIFLIFQTKNVEA ASCPNVGAVCSPFETKPCGNVKDCRCLPWGLFF-GTCINPTGSKYNMKMIEEHPNLC/...
Q9FRT9 (Van, AB052880)                                              /A ADCNGA---CSPFQMPPCGSTD-CLCIPAGLLFVGYCTYPSGLSSVAKMIDEHPNLC/...
Q9FRT8 (Vra, AB052881)                                              /A ADCNGA---CSPFEMPPCRSTD-CRCIPIALFG-GFCINPTGLSSVAKMIDEHPNLC/...
Consensus (A1b/seeds)                                                a   c g    cspf  ppc     crc p  l   g c  p g     +  dehpnlc
```

Fig. 4. Sequence alignment of reference PA1b peptides, and peptide sequences deduced from conceptual translation of the cloned genes from the four studied legume species and genotypes. First column indicates peptide identification (Swissprot accession, or isoform label, or EST accession), then the plant source, eventual gene name and gene accession (EMBL). Plant species: Psa, *Pisum sativum*; Gma, *Glycine max*; Gso, *Glycine soya*; Pvu, *Phaseolus vulgaris*; Mtr, *Medicago truncatula*; Van, *Vigna angularis*; Vra, *Vigna radiata*. First line indicates canonical (pea-type) processing of the pre-propeptide (Higgins et al. [7]). Conserved bridged cysteine residues are highlighted (1–4, 2–5, 3–6 knottin-type bridges; relative cysteine numbering within PA1b).

is presented for bean (PvuA1b041; Fig. 4), owing to the presence of two 5′-gene walking clones different from each other and from the initial degenerate PCR sequence. This indicates the presence of other A1b loci/alleles in *P. vulgaris*, which, in turn, conforms with the presence of at least three unmatched peptide hits in the aqueous methanolic fraction of this species (see Section 3.5 and Table 2).

The global structure of the genes is preserved in all the species, including *M. truncatula*, coding for a signal peptide, incorporating a 82–368 bp intron (from pea to soybean, not shown) followed by PA1b, a short linker peptide of variable size, and the better conserved PA1a peptide (preproprotein sequences shown in Fig. 4). The gene structure and sequence conservation supports the homology of the obtained sequences with the PA1b gene.

When compared with published sequences, the pea and soybean sequences correspond to new alleles with few mutations (all synonymous, as compared to some existing EST hits for the soybean sequence, Fig. 4), with the exception of one 100% nucleic acid identity with pea sequence EMBL:M81864 for our AJ574795 gene. Also, pea sequence AJ574796 codes for a variant peptide, PsA1b015, already described at the protein (but not nucleic acid) level by Higgins et al. [7]. Together with the peptide match results described in the next section, this is a strong indication that all our sequences from *P. sativum* represent slightly divergent independent genes in this species. This complex situation does not seem to prevail in soybean, whether analysed in EST sequences or in peptide extracts (Section 3.5). The new genes from *P. vulgaris* and *M. truncatula* increase the known variability of the family, leading to the determination of conserved amino acids in A1b peptides (consensus sequence in Fig. 4), as follows: the six structural cysteines are conserved, apart from the last one in MtrA1b006; the five prolines and glycines G5 and G30 (pea numbering), plus arginine R21 and leucine L27 are also conserved.

3.5. Peptide characterisation in hydrophobic peptide fractions

SDS-PAGE identified polypeptides from all plant fractions except the first methanolic extracts (not shown) and the *M. truncatula* MeOH60 fraction. The aqueous methanolic fractions, which contained the highest A1b-like activities (except in *M. truncatula*; Figs. 2 and 3, Table 1), were analysed by SDS-PAGE and Western-blotting (Fig. 5) and by MALDI–TOF mass spectrometry (Table 2). Both methods identified peptides in the 4 kDa range from all plants except from *M. truncatula*, for which only a very low level of peptides was detected by mass spectrometry. The quantity of protein determined by our Bradford assay in this MeOH60 fraction (coloured extracts) was obviously overestimated. An additional peptide group in the 6 kDa range was also detected, abundant only in pea (Fig. 5A and Table 2). SDS-PAGE gel showed a protein of around 10 kDa in the soybean MeOH60 extract, which may be the soybean hydrophobic protein, another identified storage protein [20]. Anti (pea) PA1b antibodies recognised only the 4 kDa band from pea (Fig. 5B), even when higher peptide loads were used. Assuming canonical pea-like post-translation processing (Fig. 4), we were able to match all the cloned A1b sequences with the corresponding peptides in the MeOH60 extracts, except for *M. truncatula* (Table 2). Additionally, C-terminal glycine-trimmed variants of the full A1b peptides [12] were identified from the three A1b-positive species (Table 2), and additional A1b peptides could be assigned to unmatched masses by identifying the G37 alternative trimming. Peptides of MW 3822, 3866, and 3874 are most probably other A1b isoforms in bean as we are able to identify their −57 (-glycine) counterparts. In pea, an additional PA1b isoform, not sequenced at the gDNA or cDNA level in

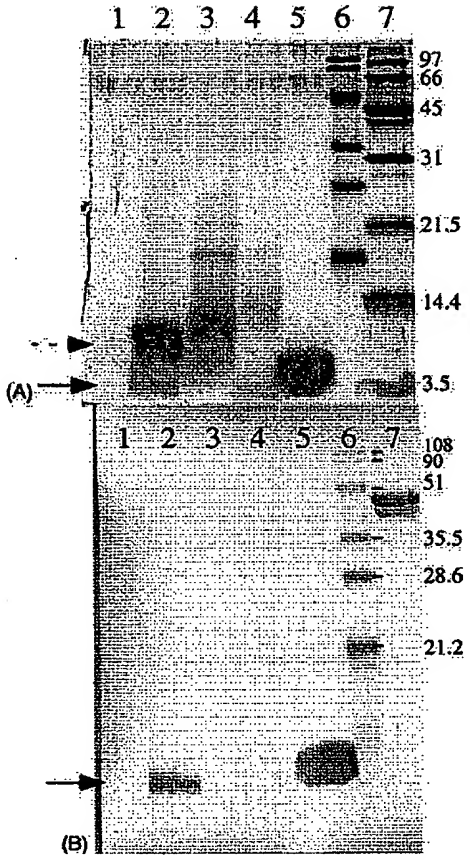


Fig. 5. SDS-PAGE electrophoresis and Western-blotting of peptides from aqueous methanolic (MeOH60) fractions from seeds of the four plant species used. Lanes (protein load): (1) *M. truncatula* (10 μg), (2) *P. sativum* (20 μg), (3) *G. max* (20 μg), (4) *P. vulgaris* (10 μg), (5) purified pea A1b, (6) low molecular weight markers (Bio-Rad LMW calibration kit plus insulin), and (7) coloured markers. (A) Coomassie G stain of electrotransferred gel, (B) Western-blot with anti (pea) PA1b antibody and alkaline phosphatase detection. Plain arrow indicates PA1b peptides (pea, soybean and bean lanes), dotted arrow points to the pea PA1a peptides (pea lane only).

this work but identified previously by full peptide sequencing [5], matched the 3758 MH+ peak in the MeOH60 extracts (Table 2). Also, pea was the only plant in which we identified the four cloned PA1a peptides, in addition to the toxic PA1b peptides, in the MeOH60 fractions (Table 2). Whether A1a was completely absent, or was in fact present in other fractions from bean and soybean extracts, has not been established. In *M. truncatula*, the possibility of variant processing of an A1b-like peptide was further assessed by fitting the observed masses to all possible C-terminal variants, in the 3–6 kDa range, from the two available nucleic sequences (Fig. 4), but without success.

In the more complex pea extracts, we analysed further the individual peptide content of the MeOH60 extracts by HPLC followed by ESI–MS, as well as peptide cross-contamination between successive fractions (Table 2). HPLC separated six groups of peptides with the following molecular masses (given with their peptide isoform perfect matches when available, excluding the minor matches with the glycine-trimmed forms, see Table 2): peak1, 3788.73 ± 0.08 (isomass to PsaA1b015? see peak 3) and 3741.43 ± 0.35 (PsaA1b011 and PsaA1b014); peak 2, 3757.68±0.60; peak 3, 3788.48±0.15 (PsaA1b015?); peak 4, 3772.53±0.08, 3842.73±0.08 and 3819.48±0.98; peak 5, 3519.78±0.29 and 3502.53±0.08; peak 6, 3789.48±0.14 (PsaA1b012). Overall, these matches, with the higher resolution of LC–MS, are a strong indication that at least 10 independent genes and peptides (loci and/or alleles) are expressed in pea, from which we identified only four by our random screening at genomic and cDNA levels. It may be noted that different genes may encode slightly different A1b peptides with identical molecular masses (isoforms 011 and 014; or the 015 isomass forms in peaks 1 and 3), which indicates a complex and probably recently divergent evolution of these homologous genes in this species.

Finally, the comparison of peptide matches in three successive fractions from pea indicated that some extracts were only marginally contaminated by major components from subsequent fractions (e.g. the methanolic fraction), while others may share major products as a result of incomplete extraction at the previous step (e.g. PA1a isoform PsaA1a012 in both aqueous methanol and acidic water extracts).

4. Discussion

A differential toxicity between susceptible and resistant weevil strains was detected on different seed extracts for each of the four plants tested. As the acute toxicity of pea seeds is due to PA1b [5], and the resistance of *S. oryzae* to this acute toxicity was shown to be a monogenic character [6], we hypothesised that a differential R/S toxicity was due to A1b-like peptides. This assumption proved to be true in three of our test panel plants, for which a total differential mortality was present (no toxicity to the R strain). This simple toxicity pattern, due to A1b only, was therefore extended from a plant within the *Viciae* (pea) to two species within the *Phaseolae* tribe (bean and soybean); it should be noted that when assayed at a lower concentration (one-fifth of the dose given in Table 1) the toxicity from these three species is displayed mainly in the MeOH60 fractions (not shown) containing most of the A1b peptides. In *M. truncatula* however, MeOH and H_2O5 fractions showed both high toxicities to the R strain and a significant extent of differential toxicity, whereas the MeOH60 extract did not display such a differential. No pea-like processed PA1b was found in this species in the expected apolar fractions therefore raising the dual, and non-exclusive, possibility that either a differential toxicity might exist in the absence of A1b peptides (methanolic fraction), or that a different homologous peptide exists in the aqueous fractions from this species (not yet analysed by mass spectrometry). The latter possibility is likely as a longer peptide form was detected in *M. truncatula* whole seed extracts by Western-blotting with an anti pea PA1a peptide (L. Quillien, unpublished results). Examination of the *M. truncatula* sequences may support this hypothesis of a variant post-translation processing (propeptide maturation), as these sequences were the only ones (i) lacking the C-terminal glycine (identified as a processing variant site by MS in all other species) and (ii) displaying a downstream proline, a residue known to inhibit the activity of endoproteolytic enzymes when present at their substrates' P1 site [21].

As for the binding competition assay, our results show that this assay was able to detect A1b-like activity in rather crude extracts from three plants, thereby correlating well with the insect bioassay. However, in *M. truncatula* methanolic extract, from which significant differential toxicity was detected, we were not able to detect either A1b-like peptides or significant ^{125}I–PA1b-binding competition. This may indicate either the presence of an inhibitor of the binding assay and of mass spectrometry ionisation (i.e. of the detection of strict A1b-like peptides), or the presence in methanolic extracts of toxic compounds differing from the A1b peptides and thus not competing for ^{125}I–PA1b binding, for which the "China" R strain was also partly resistant. Both hypotheses are plausible as (i) *M. truncatula* mass spectra were uniquely characterised by the presence of low peptide content and peak groups characteristic of polysaccharidic (hexosamine) oligomers possibly interfering with peptide ionisation and (ii) the R strain was collected in China on undefined seed stores. Further biochemical and genetic analyses of the barrel medic/weevil interaction will be needed to understand this relationship.

In contrast to the complex pattern in *M. truncatula*, the peptide load of the pea, bean and soybean toxic extracts appears to be simple. Around 70% of the matched masses in MeOH60 fractions belonged either to A1b (soybean, and probably bean) or to the sum of A1a and A1b peptides (pea); intriguingly, we were not able to identify A1a peptides from bean and soybean aqueous methanolic extracts (Table 2 and Fig. 5A). From analysing only the sequence information,

this is difficult to explain as PA1a is very similar in the three species. Different turnovers of the mature A1a peptide in the growing seeds of the Phaseoleae may account for such differences, but very limited data exist on legume seed contents in such small hydrophobic peptides, and A1a protein has actually only been reported before from pea [7,22]. Purified PA1a is not toxic to weevils (Quillien and Delobel, unpublished), and its function is still unknown. A1b was also termed leginsulin in soybean because of its ability to compete with insulin for binding with a seed 7S globulin [12], but this insulin analogy was shown to play no role in the insect-toxic syndrome since bovine insulin did not influence either toxicity or binding to the target in the weevils [11].

Another intriguing feature of our study was the absence of a signal detected from non-pea seed extracts in Western-blots with an anti-PA1b antibody, despite high sequence similarities between the targeted A1b peptides (Fig. 4). PA1b proved to be feebly antigenic, and serum titres and detection limits were weak (Fig. 5B). Western-blotting is therefore not a suitable tool for screening A1b peptides in legume seeds. Sequence comparisons and the recent structural analysis of the PA1b peptide [9] enable us to propose a potential epitope area for this antibody, with the CGTSAC sequence (pea), which forms a variable, hydrophilic and exposed loop within the PA1b structure [9].

In our combined approach, the MALDI–TOF mass spectrometry data on relatively crude extracts proved to be of critical importance for both matching genomic PCR data, thus clarifying small sequence variability due to potential PCR errors, and for having a clear picture of the peptide complexity of extracts. Apart from *M. truncatula*, in which the uncertainty persists, all the sequenced genes were shown to be expressed and producing canonical A1b peptides in seeds. Careful analysis of masses (for identification of potential C-terminal variants) should allow a systematic identification of the number of expressed genes (either loci or, at least, alleles). This feature is undoubtedly quite variable, leading to a minimal loci/allele estimate of 1–3 in soybean, of 2–4 in bean and of 5–10 in pea (from mass data). In soybean, only one tentative consensus group (TC, TIGR TC160324), corresponding to one unigene entry (NCBI UNIGENE Gma.188), was identified for leginsulin from extensive EST analysis, therefore predicting only one locus for this species. Overall, this species-specific low-level variability, together with the identification of a main genomic localisation of PA1b loci on pea chromosome 6 (EMBL: AJ276882), may indicate very recent duplication events yielding to clustered arrays of A1 genes in some species, possibly related to domestication. Ongoing genetic analysis of this loci actually estimates seven PA1 alleles in pea, six of which are polymorphic between genotypes and co-segregate at one locus on linkage group 6 (Domoney, personal communication). The intensity of acute seed insecticidal activity has been screened in pea genotypes [6] and varies considerably, but the extent to which this trait is driven by PA1b expression levels is currently not known. Potentially similar trait variability and gene structure has been recently described for the pea seed Bowman–Birk trypsin inhibitors [23].

We have chosen to retain a neutral nomenclature for identifying isoforms, with GspA1xn as a template, as used in Fig. 4 (Gsp for genus species identifiers, A1 for albumin 1 family—not leginsulin, x for potential mature peptide naming, and *n* as a three digit numbering for the database submitted sequences identifying unique loci/allelic variants at the cds nucleic sequence level). At the protein sequence level, we analysed the PA1b variability of known peptides from legume seeds, for which a tentative consensus sequence is given (Fig. 4). In the two genomic model species, comprehensive EST analysis (successive blast rounds starting with pea sequences, together with the use of the TIGR gene indices databases) led to the identification of more than 25 TC clusters with some degree of homology to PA1b in *M. truncatula* (in addition to the two new genes sequenced in our work), and four additional TCs only found in soybean. As suggested from expression analysis in pea [7] or soybean and barrel medic (http://www.tigr.org/tdb/tgi/plant.shtml), the insect-toxic A1bs are tightly regulated for expression in seeds (embryos/cotyledons) while the other genes, quite different in sequence apart from the cysteine topology and their association with the more conserved A1a peptide (not shown), are not associated with expression in seeds. It seems reasonable to hypothesise that the two genes identified here in *M. truncatula* are actually seed expressers, due both to their high sequence homology to PA1b/leginsulin (>55% identity at the peptide level) and to the lack of EST coverage for maturing seeds in this species.

At the gene level, the two-exon structure, associating a short variable intron inserted in the C-terminally conserved signal peptide (conserved alanine, Fig. 4), is a common feature of all the identified genes from this family. Interestingly, this gene structure is reminiscent of the organisation of a set of cysteine-rich peptides from mustard and *Arabidopsis* [24], which are clustered in a tight array of four successive genes in the *Arabidopsis* genome. These genes are expressed in seeds, but they are also induced by wounding [25] and by nematodes [26], therefore constituting a complex of tandem-organised defensive peptides. Finally, in the identified *M. truncatula* genes, the presence of transcription regulation signals in the 5′ sequenced regions (among which are the putative TATA boxes) was a further indication that these genes are functional, despite the lack of any detection of their products in the present work.

In conclusion, through the use of combined approaches validated on pea as the type species, we demonstrated the presence of both standard and variant A1b peptides and/or genes in two new species, one belonging to the same complex tribe as the soybean (bean, Phaseoleae) the other belonging to a sister tribe to the Vicieae (barrel medic, Trifolieae) [27]. These results indicate that the A1b family, currently unidentified in non-legume plants, is probably well represented in the Fabaceae. Extending this study, using the combined approach presented here, appears to be

feasible with the dual aim of detecting new peptides with variant biological activities and of obtaining clues to the evolutionary history of this recently discovered entomotoxic peptide family.

Acknowledgements

We are grateful to Christophe Chambon (INRA Clermont-Theix) for the MALDI–Tof/MS analysis, to Daniel Mollé (INRA, Rennes) for the ESI/MS analysis, and to Gabrielle Duport for her help in weevil mass rearing and preparation. We thank Valerie James for the English revision. We fully acknowledge financial support from INRA Plant Health department to YR, and from INRA (Departments "Santé des Plantes et Environnement" and "Technologie des Produits Végétaux") for a PhD grant to SL.

References

[1] L.M. Schoonhoven, T. Jermy, J.J.A. van Loon, Insect-Plant Biology, From Physiology to Evolution, Chapman & Hall, London, GBR, 1998, 409p.
[2] G.C. Yencho, M.B. Cohen, P.F. Byrne, Applications of tagging and mapping insect resistance loci in plants, Annu. Rev. Genet. 45 (2000) 393–422.
[3] S. Muthukrishnan, G.H. Liang, H.N. Trick, B.S. Gill, Pathogenesis-related proteins and their genes in cereals, Plant Cell Tissue Organ Cult. 64 (2001) 93–114.
[4] G.S. Khush, D.S. Brar, Genetics of resistance to insects in crop plants, Adv. Agron. 45 (1991) 223–274.
[5] B. Delobel, A.M. Grenier, J. Gueguen, E. Ferrasson, M. Mbaiguinam, Utilisation d'un Polypeptide dérivé d'une albumine PA1b de légumineuse Comme Insecticide, Patent-98/05877, Paris, 1998, pp. 1–25.
[6] A.M. Grenier, M. Mbaiguinam, B. Delobel, Genetical analysis of the ability of the rice weevil *Sitophilus oryzae* (Coleoptera, Curculionidae) to breed on split peas, Heredity 79 (1997) 15–23.
[7] T.J.V. Higgins, P.M. Chandler, P.J. Randall, D. Spencer, L.R. Beach, R.J. Blagrove, et al., Gene structure, protein structure, and regulation of the synthesis of a sulfur-rich protein in pea seeds, J. Biol. Chem. 261 (1986) 11124–11130.
[8] M. Mbaiguinam, Détermination des Causes de la résistance des Légumineuses aux Charançons des céréales du genre *Sitophilus*, PhD, Univ. Claude Bernard Lyon1, Lyon, 1996, pp. 1–164.
[9] L. Jouvensal, L. Quillien, B. Ferrasson, Y. Rahbé, J. Gueguen, F. Vovelle, PA1b, an insecticidal protein extracted from pea seeds (*Pisum sativum*): (1)H-2D-NMR study and molecular modeling, Biochemistry 42 (2003) 11915–11923.
[10] T. Yamazaki, M. Takaoka, E. Katoh, K. Hanada, M. Sakita, K. Sakata, et al., A possible physiological function and the tertiary structure of a 4-kDa peptide in legumes, Eur. J. Biochem. 270 (2003) 1269–1276.
[11] F. Gressent, I. Rahioui, Y. Rahbé, Characterisation of a high affinity binding site for the pea albumin 1b (PA1b) entomotoxin in the weevil *Sitophilus*, Eur. J. Biochem. 270 (2003) 2429–2435.
[12] Y. Watanabe, S.F. Barbashov, S. Komatsu, A.M. Hemmings, M. Miyagi, S. Tsunasawa, et al., A peptide that stimulates phosphorylation of the plant insulin-binding protein. Isolation, primary structure and cDNA cloning, Eur. J. Biochem. 224 (1994) 167–172.
[13] S.C. Ilgoutz, N. Knittel, J.M. Lin, S. Sterle, K.R. Gayler, Transcription of genes for conglutin gamma and a leginsulin-like protein in narrow-leafed lupin, Plant Molec. Biol. 34 (1997) 613–627.
[14] K. Hanada, Y. Nishiuchi, H. Hirano, Amino acid residues on the surface of soybean 4-kDa peptide involved in the interaction with its binding protein, Eur. J. Biochem. 270 (2003) 2583–2592.
[15] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254.
[16] J.P. Briand, S. Muller, M.H. Van Regenmortel, Synthetic peptides as antigens: pitfalls of conjugation methods, J. Immunol. Methods 78 (1985) 59–69.
[17] F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, et al., Current Protocols in Molecular Biology, John Wiley, New York, 1997.
[18] D. Page, Genetic Improvement of the Pea Seed Storage Protein: Heredity of its Composition and Molecular Study of the Tri Locus, QTL Candidate of the Trypsin Inhibitor Activity, PhD Thesis, Inst. Nat. Polytechnique Lorraine, 1999, 122 p.
[19] R.A. Salzman, T. Fujita, K. Zhu-Salzman, P.M. Hasegawa, R.A. Bressan, An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates, Plant Molec. Biol. Rep. 17 (1999) 11–17.
[20] F. Baud, E. Pebay-Peyroula, C. Cohen-Addad, S. Odani, M.S. Lehmann, Crystal structure of hydrophobic protein from soybean a member of a new cysteine-rich family, J. Mol. Biol. 231 (1993) 877–887.
[21] N.D. Rawlings, A.J. Barrett, MEROPS: the peptidase database, Nucleic Acids Res. 28 (2000) 323–325.
[22] J.A. Gatehouse, J. Gilroy, M.S. Hoque, R.R. Croy, Purification, properties and amino acid sequence of a low-Mr abundant seed protein from pea (*Pisum sativum* L.), Biochem. J. 225 (1985) 239–247.
[23] D. Page, G. Aubert, G. Duc, T. Welham, C. Domoney, Combinatorial variation in coding and promoter sequences of genes at the Tri locus in *Pisum sativum* accounts for variation in trypsin inhibitor activity in seeds, Mol. Genet. Genom. 267 (2002) 359–369.
[24] L.R. Ceci, N. Spoto, M. de Virgilio, R. Gallerani, The gene coding for the mustard trypsin inhibitor-2 is discontinuous and wound-inducible, FEBS Lett. 364 (1995) 179–181.
[25] F. De Leo, L.R. Ceci, L. Jouanin, R. Gallerani, Analysis of mustard trypsin inhibitor-2 gene expression in response to developmental or environmental induction, Planta 212 (2001) 710–717.
[26] I. Vercauteren, E. Van Der Schueren, M. Van Montagu, G. Gheysen, *Arabidopsis thaliana* genes expressed in the early compatible interaction with root-knot nematodes, Mol. Plant Microb. Interact. 14 (2001) 288–299.
[27] N.D. Young, J. Mudge, T.H. Ellis, Legume genomes: more than peas in a pod, Curr. Opin. Plant Biol. 6 (2003) 199–204.